

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

HO SANG SUNG, ET AL.

Application No.:

Filed:

For: **LSF COEFFICIENT VECTOR  
QUANTIZER FOR WIDEBAND  
SPEECH CODING**

Art Group:

Examiner:

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

**REQUEST FOR PRIORITY**

Sir:

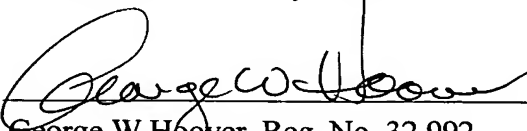
Applicant respectfully requests a convention priority for the above-captioned application, namely:

| COUNTRY | APPLICATION<br>NUMBER | DATE OF FILING |
|---------|-----------------------|----------------|
| Korea   | 2003-13606            | 5 March 2003   |

☒ A certified copy of the document is being submitted herewith.

Respectfully submitted,

Blakely, Sokoloff, Taylor & Zafman LLP



George W Hoover, Reg. No. 32,992

Dated: December 30, 2003

12400 Wilshire Boulevard, 7th Floor  
Los Angeles, CA 90025  
Telephone: (310) 207-3800



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출원 번호 : 10-2003-0013606  
Application Number

출원 년 월 일 : 2003년 03월 05일  
Date of Application MAR 05, 2003

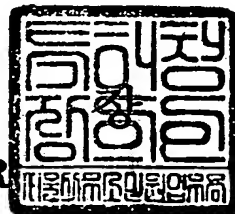
출원인 : 한국전자통신연구원  
Applicant(s) Electronics and Telecommunications Research Institut



2003      03      20  
년      월      일

특      허      청

COMMISSIONER





## 【서지사항】

|            |  |
|------------|--|
| 【서류명】      | 특허출원서  |
| 【권리구분】     | 특허   |
| 【수신처】      | 특허청장   |
| 【참조번호】     | 0001   |
| 【제출일자】     | 2003.03.05   |
| 【국제특허분류】   | G11B   |
| 【발명의 명칭】   | 광대역 음성 부호화를 위한 엘에스에프 계수 벡터 양자화기                                |
| 【발명의 영문명칭】 | Quantizer of LSF coefficient vector in wide-band speech coding |
| 【출원인】      |  |
| 【명칭】       | 한국전자통신연구원  |
| 【출원인코드】    | 3-1998-007763-8  |
| 【대리인】      |  |
| 【성명】       | 이영필  |
| 【대리인코드】    | 9-1998-000334-6  |
| 【포괄위임등록번호】 | 2001-038378-6  |
| 【대리인】      |  |
| 【성명】       | 이해영  |
| 【대리인코드】    | 9-1999-000227-4  |
| 【포괄위임등록번호】 | 2001-038396-8  |
| 【발명자】      |  |
| 【성명의 국문표기】 | 성호상  |
| 【성명의 영문표기】 | SUNG, Ho Sang  |
| 【주민등록번호】   | 710501-1822714   |
| 【우편번호】     | 305-325  |
| 【주소】       | 대전광역시 유성구 노은동 열매마을아파트 901동 401호                                |
| 【국적】       | KR   |
| 【발명자】      |  |
| 【성명의 국문표기】 | 황대환  |
| 【성명의 영문표기】 | HWANG, Dae Hwan  |
| 【주민등록번호】   | 630716-1010012   |

|            |   |
|------------|---|
| 【우편번호】     | 305-729   |
| 【주소】       | 대전광역시 유성구 전민동 청구나래아파트 103동 905호   |
| 【국적】       | KR  |
| 【발명자】      |   |
| 【성명의 국문표기】 | 강상원   |
| 【성명의 영문표기】 | KANG, Sang Won  |
| 【주민등록번호】   | 580320-1671318  |
| 【우편번호】     | 425-791   |
| 【주소】       | 경기도 안산시 사1동 한양대학교 제 3공학관 522-2 통신<br>및 신호처리 연구실                                       |
| 【국적】       | KR  |
| 【발명자】      |   |
| 【성명의 국문표기】 | 이강은   |
| 【성명의 영문표기】 | LEE, Kang Eun   |
| 【주민등록번호】   | 761016-1273218  |
| 【우편번호】     | 425-791   |
| 【주소】       | 경기도 안산시 사1동 한양대학교 제3공학관 522-2 통신 및<br>신호처리 연구실  |
| 【국적】       | KR  |
| 【심사청구】     | 청구  |
| 【취지】       | 특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규<br>정에 의한 출원심사 를 청구합니다. 대리인<br>이영필 (인) 대리인<br>이해영 (인) |
| 【수수료】      |   |
| 【기본출원료】    | 20 면 29,000 원   |
| 【가산출원료】    | 1 면 1,000 원   |
| 【우선권주장료】   | 0 건 0 원   |
| 【심사청구료】    | 6 항 301,000 원   |
| 【합계】       | 331,000 원   |
| 【감면사유】     | 정부출연연구기관  |
| 【감면후 수수료】  | 165,500 원   |
| 【기술이전】     |   |
| 【기술양도】     | 희망  |
| 【실시권 허여】   | 희망  |



1020030013606

출력 일자: 2003/3/26

【기술지도】

희망

【첨부서류】

1. 요약서·명세서(도면)\_1통

**【요약서】****【요약】**

본 발명은 음성 부호화에 관련된 것으로서 특히, 광대역 음성 부호화의 효율 및 성능에 가장 많은 영향을 미치는 LSF(Line Spectral Frequency) 계수 벡터의 양자화기에 관한 것이다. 본 발명은 기존 음성 부호화기의 LSF 계수 양자화기의 예측 구조내에 비구조적 벡터 양자화기와 격자 양자화기(피라미드 벡터 양자화기)를 상호 직렬로 연결한 형태로 대체 사용하여 양자화기에 필요로 하는 메모리 양과 탐색 시간을 줄이고, 예측 구조와 비예측 구조를 병렬로 사용함으로써 안전한 양자화와 에러 전달 현상을 줄여 할당되는 비트와 SD(Spectral Distortion)를 개선시킬 수 있는 효율적인 LSF 양자화기를 제공한다. 본 발명은 또한 비구조적 벡터 양자화를 피라미드 벡터 양자화 이전에 실시함으로써 피라미드 벡터 양자화기에 적합한 Laplacian 모델로 입력값을 변환할 수 있으며, 피라미드 벡터 양자화기의 적은 계산량을 이용하여 두 개 직렬 양자화기간의 공동 최적 벡터를 결정함으로써 성능이 우수한 양자화기를 제공한다. 아울러 예측 구조와 비예측 구조의 채택으로 예측 구조에 비적합한 외좌점을 올바르게 양자화 할 수 있는 LSF 양자화기를 제공한다.

**【대표도】**

도 3



**【명세서】**

**【발명의 명칭】**

광대역 음성 부호화를 위한 엘에스에프 계수 벡터 양자화기{Quantizer of LSF coefficient vector in wide band speech coding}

**【도면의 간단한 설명】**

도 1은 3GPP 규격에 의한 광대역 음성 부호화기에서 사용된 선형 예측 계수 양자화기의 구성도이다.

도 2는 PCPVQ을 이용한 양자화기의 구성도이다.

도 3은 본 발명의 최적의 실시예의 구성도이다.

<도면의 주요 부분에 대한 부호 설명>

30 : 예측 구조    31 : 비예측 구조

VQ1 : 제1 벡터 양자화기    PVQ1 : 제1 피라미드 벡터 양자화기

VQ2 : 제2 벡터 양자화기    PVQ2 : 제2 피라미드 벡터 양자화기

**【발명의 상세한 설명】**

**【발명의 목적】**

**【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<8>      본 발명은 음성 부호화에 관련된 것으로서 특히, 광대역 음성 부호화의 효율 및 성능에 가장 많은 영향을 미치는 LSF(Line Spectral Frequency) 계수 벡터의 양자화기에 관한 것이다.



- <9> 디지털 시대의 도래로 현재 거의 모든 통신 수단의 송수신 신호 처리 방식은 기존의 아날로그 방식을 탈피하여 디지털 방식으로 전이되었으며 보다 더욱 향상된 디지털 처리 방식 기술이 다수 출현하고 있다. 송수신 신호는 결국 영상 신호와 음성 신호이며 이들 신호를 디지털 방식으로 효율적인 전송을 위해, 전송 과정에서의 송수신 시스템의 부하를 줄이고 수신단에서 양질의 아날로그 신호로 복원하기 위해서는 상기 신호의 고품질·고효율 압축(compression, 코딩)이 필수적이다. 따라서 상기 디지털 처리 방식 기술은 두 신호를 보다 고품질의 효율적으로 압축하는 방식에 큰 비중을 두고 있다.
- <10> 이 중에서 광대역 음성 신호의 코딩은 전송시의 시스템의 부하를 줄이고, 수신단에서 고품질의 음성을 재현하기 위해서 음성 신호의 단구간 상관도를 나타내는 LSF(Line Spectral Frequency) 계수를 어떻게 효율적으로 양자화하는 것이 가장 중요한 관건이 되는데 음성 신호는 영상 신호와는 달리 단구간 상관의 정도가 떨어지기 때문이다. 따라서 단구간 상관 정도의 정확한 산출은 음성 신호의 효율적 코딩에 있어서 매우 중요하다.
- <11> 현재까지 알려진 광대역 음성 신호의 코딩 방식의 대부분은 음성의 스펙트럴 포락선(spectral envelope)을 분석하여 음성을 파라미터적으로 표현하여 이루어진다. 이 때 스펙트럴 포락선을 파라미터적으로 표현함에 있어서는 LPC(Linear Prediction Coding) 파라미터를 사용하게 되며 LPC은 단기간 선형 예측 코딩(short term linear prediction coding)이라고도 불린다.
- <12> 광대역 음성 신호의 코딩 및 디코딩은 LPC 파라미터를 양자화(quantization)하여 수신단으로 전송하고, 수신단에서는 양자화된 LPC 파라미터를 이용하여 스펙트럴 포락선을 재구성하는 과정을 수반한다.





- <13> 이 때, LPC 파라미터의 양자화는 구체적으로는 LPC 필터(선형 예측 필터)를 통하여 이루어지는데 우선 이 필터의 최적 선형 예측 계수를 구하게 된다. 최적 선형 예측 계수는 음성 신호를 프레임 단위로 나누어 각 프레임 별로 예측 오차를 최소화시키는 방향으로 구해진다.
- <14> 기존에 제시된 선형 예측 필터의 일례로서 AMR-WB(G.722.2) 음성 부호화기의 선형 예측 필터가 있는데 이 필터는 16차(16th-order) all-pole 필터이다. 이 때 pole에 대한 선형 예측 계수들의 양자화를 위하여 다량의 비트(bit)수가 필요하다. 예를 들어 CDMA 이동통신 시스템에 사용되는 음성 코딩 방식인 IS-96A QCELP(Q Code Excited Linear Prediction)은 코딩에 필요한 전체 비트의 25% 정도를 선형 예측 계수의 양자화에 할당하고 있으며, AMR-WB 음성 부호화기는 최저 9.6%에서 최고 27.3%을 할당하고 있다.
- <15> 그 간 양자화에 대한 다양한 방식이 제시되었는데 주로 선형 예측 계수를 직접 양자화하는 방식을 채용하였다. 그러나 선형 예측 계수를 직접 양자화하는 경우에는 필터의 특성이 계수의 양자화 오차에 매우 민감하며 양자화 후 필터의 안정성이 보장되지 않는 문제점이 있었다.
- <16> 이 문제점을 극복하기 위해 선형 예측 계수를 다른 표현으로 변환하여 이를 양자화시키는 기법이 개발되었다. 이 기법에서는 선형 예측 계수를 수학적으로 등가인 반사계수(reflection coefficient) 또는 LSF로 변환하여 양자화를 시행하는데, 특히 LSF은 어휘에서 의미하는 바와 같이 음성의 주파수 특성을 반영하는 것으로서 음성의 특성을 가장 잘 반영시킨다고 알려져 있으며 따라서 최근의 양자화는 대부분 선형 예측 계수를 LSF로 변환시켜 이루어지게 된다.

<17> LSF 양자화 기법은 양자화의 효율화를 위해 프레임간 상관 관계(단구간 상관도)를 이용하는데 즉, 현재 프레임의 LSF를 직접 양자화하지 않고 과거 프레임의 LSF 정보로부터 현재 프레임의 LSF를 예측한 후 이 예측의 오차를 양자화하는 기법이다. 예측의 방식으로 AR(Auto Regressive) 방식과 MA(Moving Average) 방식을 사용하게 되는데, 전자는 예측 성능이 우수한 반면에 계수 전달 오류의 영향이 수신측에 계속 전달되는 단점이 있으며 후자는 전자에 비해 예측 성능은 떨어지지만 전달 오류의 영향은 제한되는 장점이 있다. 따라서 무선 통신 등과 같이 전달 오류가 많이 발생하는 환경에서는 MA 방식을 이용하여 예측이 실시되고 있다.

<18> 일반적으로 벡터 양자화 과정에서 전체 벡터를 양자화함은 양자화를 위한 부호표(code book)의 크기가 너무 커지고 후보 벡터의 검색 시간도 많이 소요되므로 전체 벡터를 다수개의 하위 벡터로 분할하여(split) 각각 독립적으로 양자화할 필요성이 제기된다. 이를 위해 SVQ(Split Vector Quantization)가 제안되었으나 SVQ를 적용하여 양자화하더라도 여전히 부호표의 저장을 위한 많은 메모리와 계산량을 요구하기 때문에 분할의 효과는 미미하며, split의 수를 증가시키면 프레임간 상관 관계가 떨어지게 되어 양자화 성능이 저하되는 문제점이 있다.

<19> 벡터 양자화의 효율성 제고를 위한 또 다른 기법으로 다단계(multi stage)로 양자화를 구비하여 이전 단계의 양자화기에서 발생한 양자화 에러를 다음 단계의 양자화기에서 양자화 하는 기법이 제시되었다. 하지만 광대역과 같이 많은 비트를 할당한 경우 메모리와 계산량에 대한 부담이 여전히 있게 된다.

<20> 도 1에 3GPP 규격에 의한 S-MSVQ(Split-Multi Stage Vector Quantization)구조로 이루어진 광대역 음성 부호화기에서 사용된 선형 예측 계수 양자화기의 구성도가 제시되



어 있는데 이는 SVQ(Split Vector Quantization)와 다단계(multi stage) 개념이 동시에 반영된 양자화기이다. 이 양자화기에 대한 동작을 간단히 설명하면 다음과 같다.

<21> 본 양자화기에 의한 LSF 양자화는 DC 성분을 제거한 후에 16차의 LSF 계수와 예측기로 예측한 벡터와의 차이값인 16차의 예측 에러벡터를 각각 9차 및 7차의 서브벡터(dim9, dim7)로 분할(split)하여 벡터양자화를 수행한 다음, 각각의 양자화 에러벡터를 9차는 세 개의 3차 서브벡터(3개의 dim3)로 다시 분할(split)하고 7차는 3차 및 4차의 서브벡터로 분할(split)하여 이루어지게 된다.

<22> 이러한 S-MSVQ 구조는 46 bits가 할당된 LSF 계수 양자화에 요구되는 메모리와 부호표 탐색 시간을 감소시키기 위한 구조로 전체 벡터에 대한 양자화보다는 훨씬 적은 메모리와 부호표 탐색 계산량을 지나 위에서 언급한 바와 같이 여전히 많은 메모리량( $2^8 \times 9 + 2^8 \times 7 + 2^6 \times 3 + 2^7 \times 3 + 2^7 \times 3 + 2^5 \times 3 + 2^5 \times 4$ )과 부호표 탐색의 복잡성에 의한 많은 계산량이 요구된다.

<23> 벡터 양자화기는 크게 비구조적 양자화기(비격자 양자화기)와 격자 양자화기로 대별된다. 비구조적 양자화기는 부호표를 저장해야 하는 반면에 격자 양자화기는 부호표의 인덱스만을 저장하기 때문에 부호표의 메모리 용량의 측면에 있어서 비구조적 양자화기보다 장점이 있다.

<24> 격자 양자화기는 다시 균일 격자 양자화기와 의사(pseudo) 균일 격자 양자화기로 구별되며 또는 구형 격자 양자화기와 피라미드형 격자 양자화기(Pyramid Vector Quantizer, PVQ)로 구별되기도 하는데 양자화 품질, 효율 등 여러 측면을 고려할 때 주로 피라미드 격자 양자화기를 이용한다.

- <25> PVQ은 Thomas R. Fischer : "A Pyramid Vector Quantizer", IEEE Transactions on Information Theory Vol.IT-32, pp568~583, 4. JUL. 1986의 논문에 제시되어 있다.
- <26> PVQ(Pyramid Vector Quantizer)는 L 차원의 피라미드 표면에 분포하는 격자 점을 기본으로 양자화를 하여 부호표의 저장을 위한 메모리를 필요치 않으며, 벡터 차수가 증가할수록 부호화의 복잡도가 선형적으로 증가하는 특성을 가지고 있으므로 양자화시에 적은 계산량으로 벡터 전체를 양자화 할 수 있다. 특히 라플라시안 분포를 띠는 소스에 대한 PVQ의 성능은 입력 벡터 차수가 클 경우에는 엔트로피(entropy) 제한 스칼라 양자화와 거의 대등한 성능을 나타낸다.
- <27> 양자화기의 입력 벡터가 라플라시안 분포를 가질 때 최적 코드워드(codeword)는 단일 피라미드 상에 설계할 수 있다. 상기 논문에 제시된 PVQ 부호화 과정을 간단히 언급하면 다음과 같다.
- <28> 1 단계 : 입력 코드워드에 대해서 프로젝션(projection)하여 피라미드 표면의 가장 근접한 코드워드를 선택한다.
- <29> 2 단계 : 피라미드 표면으로 프로젝션 된 코드워드를 정규화 된 피라미드에 존재하도록 스케일(scale)한다.
- <30> 3 단계 : 정규화 된 피라미드위로 존재하게 된 코드워드와 가장 근접한 정수 값을 갖는 코드워드를 찾아 선택한다.
- <31> 4 단계 : 피라미드 표면의 격자 점으로 표현된 코드워드를 원래 크기로 스케일링하여 입력 코드워드의 양자화된 벡터를 구한다.

- <32> PVQ는 특히 입력 벡터의 차수가 충분히 클 때 우수한 성능을 나타낸다. 입력 벡터의 차수가 20이상일 때, 소스의 norm값이 일정한 값에 가까운 값을 갖게 된다. 그러나 벡터 차수가 20 이하일 때에는 norm값의 분산이 커져서 일정한 값으로 볼 수 없는 단점을 가진다. 그렇기 때문에 단일 피라미드를 사용하여 양자화 할 때 많은 오차가 발생하게 된다. 이러한 단점을 보완하기 위하여 상기 논문에 제시된 바와 같이 PCPVQ(Product PVQ)를 사용한다. 도 2에 PCPVQ을 이용한 양자화기의 블록도가 제시되어 있다. 이의 동작에 대한 상세한 설명은 위 논문에 자세히 나와 있으므로 약하기로 한다.
- <33> PCPVQ는 입력 벡터를 정규화 한 후 단일 피라미드를 사용하여 양자화하고, 정규화 요소값을 사용하여 양자화된 피라미드를 인덱싱(indexing)하는 방법이다. 이렇게 함으로써 정규화 요소만큼의 피라미드를 사용하는 효과를 얻을 수 있다.
- <34> 그러나 피라미드 격자 양자화기는 라플라시안(Laplacian) 분포를 띠는 소스의 처리에 적합하지만 격자 양자화기만을 사용하여 양자화를 행하는 경우에는 소스가 격자 양자화기가 지원하지 않는 분포를 가질시에 양자화 성능이 떨어지게 되는 문제점이 있다. 즉, 피라미드 격자 양자화기를 예를 들어 보면 입력으로 이용되는 예측값이 제거된 LSF 벡터는 비록 라플라시안 분포를 이룬다고는 하나 라플라시안 분포에 정확히 일치하지 않는 외좌점(outlier)의 수가 비교적 많기 때문에 양자화 성능의 저하를 피할 수 없다.
- 【발명이 이루고자 하는 기술적 과제】**
- <35> 따라서 본 발명은 상기한 문제점을 해결하기 위해 창안된 것으로, 본 발명의 목적 및 이루고자 하는 기술적 과제는 양자화에 소요되는 메모리 용량과 계산량에 대한 부담을 줄이고 격자 양자화기만을 사용하여 발생하는 양자화 성능의 저하를 방지할 수 있는 광대역 음성 부호화를 위한 LSF 계수 벡터 양자화기를 제공함에 있다.

## 【발명의 구성 및 작용】

<36>        상기의 목적 및 과제를 달성하기 위해 본 발명이 제공하는 광대역 음성 부호화를 위한 LSF 계수 양자화기는 LSF 계수 벡터를 비구조적으로 양자화하여 양자화 대상 후보 벡터를 산출하는 제1 벡터 양자화기, 상기 LSF 계수 벡터의 예측된 LSF 벡터를 산출하는 예측기, 상기 후보 벡터를 상기 예측된 LSF 벡터를 참조하여 격자 양자화시켜 상기 LSF 계수 벡터의 예측 양자화 최종 벡터를 산출하는 제1 격자 양자화기를 구비한 예측 구조 양자화기; 상기 LSF 계수 벡터를 비구조적으로 양자화하여 양자화 대상 후보 벡터를 산출하는 제2 벡터 양자화기, 상기 후보 벡터를 격자 양자화하여 상기 LSF 계수 벡터의 비예측 양자화 최종 벡터를 산출하는 제2 격자 양자화기를 구비한 비예측 구조 양자화기; 및 상기 예측 양자화 최종 벡터와 비예측 양자화 최종 벡터 중 상기 LSF 계수 벡터와의 차가 작은 것을 상기 LSF 계수 벡터의 최종 양자화 벡터로 결정하는 스위칭부를 포함함을 그 특징으로 한다.

<37>        상기 예측 구조 양자화기와 비예측 구조 양자화기는 병렬 연결되어 상기 LSF 계수 벡터를 양자화함이 바람직하다. 상기 제1 벡터 양자화기와 제1 격자 양자화기는 서로 직렬 연결되어 상기 LSF 계수 벡터를 양자화함이 바람직하다. 상기 제2 벡터 양자화기와 제2 격자 양자화기는 서로 직렬 연결되어 상기 LSF 계수 벡터를 양자화함이 바람직하다. 상기 제1 격자 양자화기는 피라미드 벡터 양자화기임이 바람직하다. 상기 제2 격자 양자화기는 피라미드 벡터 양자화기임이 바람직하다.

<38>        이하 본 발명의 구성 및 작용을 본 발명의 최적의 실시예에 근거, 첨부 도면을 참조하여 상세히 설명하되 도면의 구성요소들에 참조번호를 부여함에 있어서 동일 구성요

소에 대해서는 비록 다른 도면상에 있더라도 동일 참조번호를 부여하였으며 당해 도면에 대한 설명시 필요한 경우 다른 도면의 구성요소를 인용할 수 있음을 미리 밝혀둔다.

<39> 도 3은 본 발명이 제공하는 LSF 양자화기의 최적의 실시예의 구성도이다.

<40> 본 발명은 예측 구조(30)와 비예측 구조(31)를 병렬로 연결하여 동시에 예측 방식과 비예측 방식으로 LSF 계수 벡터( $f$ )를 양자화하며, 두 방식으로 양자화된 벡터( $\hat{f}_1, \hat{f}_2$ ) 중 하나의 벡터( $\hat{f}_{fin}$ )만 선택하는 safety-net 구조를 취하고 있다. 또한 각 구조(30,31)는 비구조적 벡터 양자화기(VQ1, VQ2)와 피라미드 벡터 양자화기(PVQ1, PVQ2)를 직렬로 결합한 형태의 다단계(multi stage) 양자화 구조로 이루어져 있다.

<41> 우선 예측 구조(30)를 이용한 양자화에 대해 설명한다.

<42> 첫 번째 단계의 양자화기, 즉, 제1 벡터 양자화기(VQ1)는 비구조적 벡터 양자화기로서 벡터 양자화를 수행하는 부분이다. 벡터 양자화를 통해 부호표(code book)에서 양자화 후보 벡터(candidate vector)를 선택하게 되는데, 입력 LSF 계수 벡터( $f$ )에서 DC값(LSF\_mean\_vector)을 제거한 벡터( $f'$ )를 구한 후,  $f'$ 와 예측기(predictor)가 산출하는 LSF 계수 벡터의 예측된 LSF 벡터( $\tilde{f}'$ )와의 에러 벡터( $r$ )를 벡터 양자화시켜 양자화된 에러 벡터( $\hat{r}_1$ )를 산출하며 이것이 후보 벡터가 된다.

<43> 제1 벡터 양자화기(VQ1)는 단구간의 상관도가 떨어지지 않도록 에러 벡터( $r$ ) 전체에 대하여 양자화 한다. 전체 벡터를 양자화 하므로 부호표의 크기를 고려하지 않을 수 없는데 본 발명에서는 전체 할당 비트 중에서 1/7 이하의 비트를 벡터 양자화에 할당함



으로써 벡터 양자화에서 소모되는 부호표 메모리와 탐색 시간의 부담을 줄일 수 있게 하였다.

<44> 두 번째 단계의 양자화기, 즉, 제1 피라미드 벡터 양자화기(PVQ1)는 격자 양자화기로서 후보 벡터를 예측된 LSF 벡터( $\tilde{f}'$ )를 참조하여 격자 양자화시켜 LSF 계수 벡터의 예측 양자화 벡터 즉, 예측 구조(30)를 통한 LSF 계수 벡터( $f$ )의 양자화 벡터( $\hat{f}_1$ )을 산출하는 부분이다. 이를 위해 우선 에러 벡터( $r$ )의 양자화된 벡터( $\hat{r}_1$ )와 원래 에러 벡터( $r$ )간의 차벡터( $e$ )를 양자화시킨다.

<45> 단일 피라미드를 사용하는 피라미드 벡터 양자화기는 입력 벡터 차수가 충분히 클 때 우수한 성능을 나타내는데 이 때 입력 벡터 차수가 충분히 크다는 것은 차수가 20 이상 일 때를 말한다. 하지만 광대역용 음성 부호화기와 같이 차수가 20 이상이 되지 못할 경우 피라미드의 크기를 나타내는 벡터의 norm의 분산 값이 커져 양자화 오차가 커지게 된다. 이러한 단점을 보완하기 위하여 PCPVQ가 상기 논문에 의해 제안되었다. 광대역용 음성 부호화기에서 선형 예측 계수의 차수는 16차이므로 본 발명에서는 이 때 제1 피라미드 벡터 양자화기(PVQ1)로 적산 코드 PVQ(Product Code Pyramid Vector Quantizer, PCPVQ)를 사용한다. 아래에서 언급될 제2 피라미드 벡터 양자화기(PVQ2)도 PCPVQ를 사용한다.

<46> PCPVQ는 양자화기 입력 벡터를 정규화 한 후 단일 피라미드를 사용하여 양자화하고, 정규화 요소 값을 사용하여 양자화된 피라미드 크기를 나타내는 방식이다. 결국 단일 피라미드가 아닌 정규화 요소 값만큼의 피라미드로 양자화 하는 것과 같은 효과를 얻을 수 있다.





- <47> 제1 피라미드 벡터 양자화기(PVQ1)는 16개의 차벡터(e)를 입력받아, 각 차벡터에 대해 PVQ를 각각 수행한다. 이를 수행하는데 필요한 계산량은 피라미드 벡터 양자화기의 매우 적은 계산량을 고려 할 때 그리 큰 문제는 되지 않는다. 결국 두 개 양자화기 간 (VQ1,PVQ1)의 공동 최적 벡터(joint optimization vector)를 결정함으로써 성능이 우수한 양자화가 실시된다.
- <48> 본 발명에서 제안하는 예측 구조(30) 양자화기의 동작을 보다 상세히 설명한다.
- <49> LSF 계수 벡터(f)는 예측 구조(30)와 비예측 구조(31)로 각각 입력된다. LSF 계수 벡터(f)로부터 평균적으로 분포하는 LSF값(DC값, LSF\_mean\_vector)을 제거하여 평균값이 제거된 LSF 벡터(f')를 획득한다. 이는 LSF 벡터(f)를 부호표의 i번째 코드 워드로 표현하는 과정이다.
- <50> 【수학식 1】  $f' = f - \text{LSF\_mean\_vector}$
- <51> 다음으로 예측기(predictor)가 산출하는 LSF 계수 벡터의 예측된 LSF 벡터( $\tilde{f}'$ )와의 에러 벡터(r)를 구하게 된다.
- <52> 【수학식 2】  $r = f' - \tilde{f}'$
- <53> r은 수학식 2에서 보는 바와 같이 평균값이 제거된 LSF 벡터(f')에서 예측 벡터  $\tilde{f}'$ 이 제거된 에러 벡터를 나타낸다.
- <54> 제1 벡터 양자화기(VQ1)는 r을 벡터 양자화한  $\hat{r}_1$ 을 산출하게 되며 이는 위에서 언급한 후보 벡터이다.  $\hat{r}_1$ 은 두 번째 단계 양자화기인 제1 피라미드 벡터 양자화기(PVQ1)에 의한 격자 양자화 처리를 위하여 PVQ에 최적인 라플라시안 분포에 접근되도록 조정된 벡터(e)로 변환된다. 이는 다음 수학식 3에 의해 이루어진다.

<55> 【수학식 3】  $e = r - \text{hat}(r_1)$

<56> 즉,  $e$ 은 원래 에러 벡터( $r$ )와  $\text{hat}(r_1)$ 의 벡터 양자화된 에러 벡터( $\text{hat}(r_1)$ )의 차이이며 라플라시안 분포에 근접한 벡터이다.

<57>  $e$ 은 제1 피라미드 벡터 양자화기(PVQ1)에 의해 양자화되어 PVQ 양자화된 차벡터( $\text{hat}(e)$ )로 변환되며,  $\text{hat}(e)$ 이 후보 벡터( $\text{hat}(r_1)$ )와 합해져 상기  $r$ 의 최종 양자화 벡터( $\text{hat}(r)$ )가 산출된다.  $f'$ 의 양자화 벡터( $\text{hat}(f')$ )는  $\text{hat}(r)$ 와  $\text{hat}(f')$ 을 합하여 산출되며, LSF 계수 벡터( $f$ )의 예측 구조(30)를 통한 최종 양자화 벡터( $\text{hat}(f_1)$ )는  $\text{hat}(f')$ 와  $\text{LSF\_mean\_vector}$ 이 합해져 산출되게 된다.

<58> 비예측 구조(31)에서 실시되는 양자화는 예측의 수행이 없으며, LSF 벡터( $f$ )에서 평균 LSF값(DC값,  $s\_snet\_LSF\_mean\_vector$ )이 제거된 LSF 벡터( $r'$ )를 제2 벡터 양자화기(VQ2)와 제2 피라미드 벡터 양자화기(PVQ2)를 통해 예측 구조(30)에서와 동일한 방식으로 양자화된 벡터( $\text{hat}(r')$ )를 획득한 후 이 벡터에  $s\_snet\_LSF\_mean\_vector$ 을 합하여 LSF 벡터( $f$ )의 비예측 구조(31)의 최종 양자화 벡터( $\text{hat}(f_2)$ )를 획득하게 된다. 이 때 제2 벡터 양자화기(VQ2)와 제2 피라미드 벡터 양자화기(PVQ2)는 각각 예측 구조(30)의 제1 벡터 양자화기(VQ1)와 제1 피라미드 벡터 양자화기(PVQ1)에 대응되는 것이며,  $\text{hat}(r_1')$ ,  $e'$ ,  $\text{hat}(e')$ 은 각각 예측 구조(30)의  $\text{hat}(r_1)$ ,  $e$ ,  $\text{hat}(e)$ 에 대응하는 값들이다.

<59> 스위칭부(32)는 예측 양자화 벡터( $\text{hat}(f_1)$ )와 비예측 양자화 벡터( $\text{hat}(f_2)$ ) 중 택 일하여 LSF 계수 벡터( $f$ )의 최종 양자화 벡터( $\text{hat}(f_{fin})$ )를 결정하는데,  $\text{hat}(f_1)$ 와

$\hat{f}_2$  중 상기 LSF 계수 벡터( $f$ )와의 차가 작은 것이 최종 양자화 벡터( $\hat{f}_{fin}$ )로 택일된다.

<60> 표 1~표 3에 AMR-WB의 LPC 양자화기로 사용되는 S-MSVQ(split and multy-stage vector quantization) 방식, PVQ 방식 및 본 발명에서 제시한 방식들 간의 성능, 계산량 및 부호표 저장을 위한 메모리 용량을 비교한 도표를 각각 나타내었다. 계산량은 WMOPS(Weighted Million Operation Per Second)를 사용하여 측정하였으며 성능은 SD(spectral distortion)를 사용하여 측정하였다.

<61> 아래 표에 제시된 바와 같이 AMR-WB S-MSVQ 보다 평균 SD 값을 0.1dB 정도 개선하였고 3dB에서 5dB사이의 외좌점이 0.001% 감소하였으며 PVQ보다는 평균 SD를 0.25dB 개선 하였고 3dB에서 5dB사이의 외좌점이 0.2%, 5dB이상의 외좌점이 0.005% 감소하여 세 개의 양자화 구조 중 가장 좋은 성능을 나타내었다.

<62> 계산량과 메모리는 AMR-WB 보다 각각 17% 및 51% 감소하였다.

<63> 【표 1】

|               | AMR-WB S-MSVQ | PVQ   | 본 발명  |
|---------------|---------------|-------|-------|
| 평균 SD[dB]     | 0.842         | 0.992 | 0.745 |
| 3 dB ~ 5dB[%] | 0.013         | 0.220 | 0.012 |
| 5dB 이상[%]     | 0             | 0.005 | 0     |

<64> 【표 2】

|       | AMR-WB S-MSVQ | PVQ    | 본 발명   |
|-------|---------------|--------|--------|
| WMOPS | 1.6814        | 0.0709 | 1.3988 |

## &lt;65&gt; 【표 3】

|      | AMR-WB S-MSVQ | PVQ | 본 발명 |
|------|---------------|-----|------|
| Word | 6880          | 336 | 3343 |

<66> 이제까지 본 발명에 대하여 그 바람직한 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다.

<67> 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

## 【발명의 효과】

<68> 본 발명은 기존 음성 부호화기의 LSF 계수 양자화기를 새로운 형태의 비구조적 벡터 양자화기와 격자 양자화기를 직렬로 연결한 형태로 대체 사용하여 양자화기에 필요로 하는 메모리 양과 탐색 시간을 줄이고, 예측 구조와 비예측 구조를 병렬로 사용함으로써 안전한 양자화와 에러 전달 현상을 줄여 할당되는 비트와 SD를 개선시킬 수 있는 효율적인 LSF 양자화기를 제공한다.

<69> 또한, 비구조적 벡터 양자화를 피라미드 벡터 양자화 이전에 실시함으로써 피라미드 벡터 양자화기에 적합한 Laplacian 모델로 입력값을 변환할 수 있으며, 피라미드 벡터 양자화기의 적은 계산량을 이용하여 두 개 직렬 양자화기간의 공동 최적 벡터를 결정

함으로써 성능이 우수한 양자화기를 제공한다. 아울러 예측 구조와 비예측 구조의 채용으로 예측 구조에 비적합한 외좌점을 올바르게 양자화 할 수 있는 이점도 제공한다.

**【특허청구범위】****【청구항 1】**

LSF 계수 벡터를 비구조적으로 양자화하여 양자화 대상 후보 벡터를 산출하는 제1 벡터 양자화기, 상기 LSF 계수 벡터의 예측된 LSF 벡터를 산출하는 예측기, 상기 후보 벡터를 상기 예측된 LSF 벡터를 참조하여 격자 양자화시켜 상기 LSF 계수 벡터의 예측 양자화 최종 벡터를 산출하는 제1 격자 양자화기를 구비한 예측 구조 양자화기;

상기 LSF 계수 벡터를 비구조적으로 양자화하여 양자화 대상 후보 벡터를 산출하는 제2 벡터 양자화기, 상기 후보 벡터를 격자 양자화하여 상기 LSF 계수 벡터의 비예측 양자화 최종 벡터를 산출하는 제2 격자 양자화기를 구비한 비예측 구조 양자화기; 및

상기 예측 양자화 최종 벡터와 비예측 양자화 최종 벡터 중 상기 LSF 계수 벡터와의 차가 작은 것을 상기 LSF 계수 벡터의 최종 양자화 벡터로 결정하는 스위칭부를 포함함을 특징으로 하는 광대역 음성 부호화를 위한 LSF 계수 벡터 양자화기.

**【청구항 2】**

제 1 항에 있어서,

상기 예측 구조 양자화기와 비예측 구조 양자화기를 병렬 연결시켜 상기 LSF 계수 벡터를 양자화함을 특징으로 하는 광대역 음성 부호화를 위한 LSF 계수 벡터 양자화기.

**【청구항 3】**

제 1 항 또는 제 2 항에 있어서, 상기 제1 벡터 양자화기와 제1 격자 양자화기는 서로 직렬 연결되어 상기 LSF 계수 벡터를 양자화함을 특징으로 하는 광대역 음성 부호화를 위한 LSF 계수 벡터 양자화기.

**【청구항 4】**

제 1 항 또는 제 2 항에 있어서, 상기 제2 벡터 양자화기와 제2 격자 양자화기는 서로 직렬 연결되어 상기 LSF 계수 벡터를 양자화함을 특징으로 하는 광대역 음성 부호화를 위한 LSF 계수 벡터 양자화기.

**【청구항 5】**

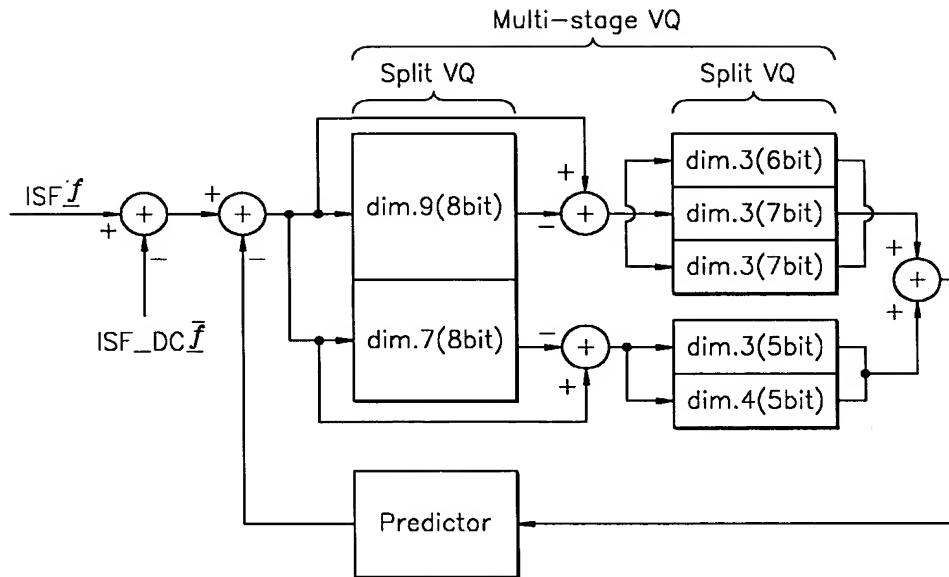
제 1 항에 있어서, 상기 제1 격자 양자화기는 피라미드 벡터 양자화기임을 특징으로 하는 광대역 음성 부호화를 위한 LSF 계수 벡터 양자화기.

**【청구항 6】**

제 1 항에 있어서, 상기 제2 격자 양자화기는 피라미드 벡터 양자화기임을 특징으로 하는 광대역 음성 부호화를 위한 LSF 계수 벡터 양자화기.

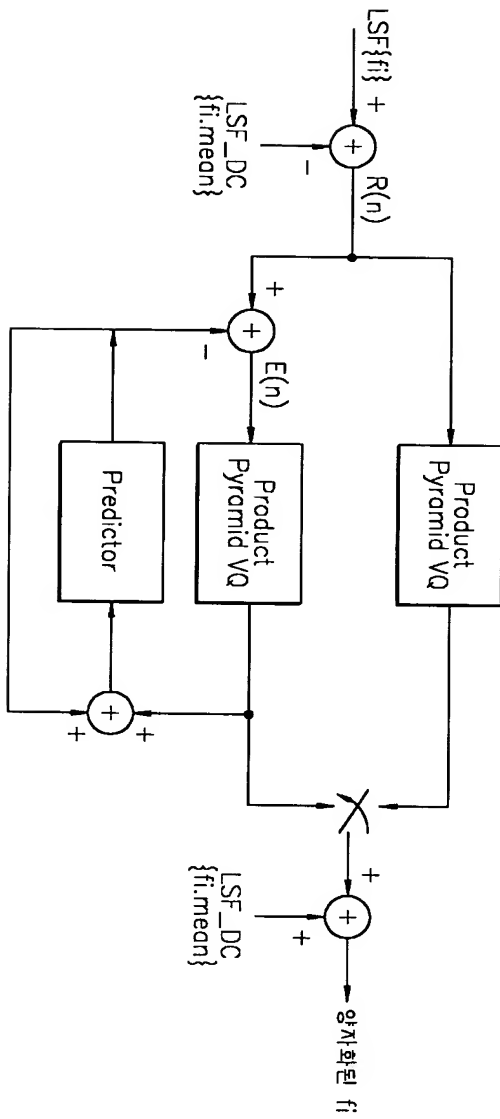
## 【도면】

【도 1】





【도 2】



【도 3】

